OIPE

#3

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/836,544A

DATE: 12/06/2001
TIME: 12:15:40

Input Set : **A:\11-881.app**
Output Set: **N:\CRF3\12062001\I836544A.raw**

ENTERED

```
 3 <110> APPLICANT: The General Hospital Corporation
 5 <120> TITLE OF INVENTION: Rapid Immunoselection Cloning Method
 7 <130> FILE REFERENCE: 11-88L
 9 <140> CURRENT APPLICATION NUMBER: US 09/836,544A
10 <141> CURRENT FILING DATE: 2001-04-17
12 <150> PRIOR APPLICATION NUMBER: US 07/983,647
13 <151> PRIOR FILING DATE: 1992-12-01
15 <150> PRIOR APPLICATION NUMBER: US 07/553,759
16 <151> PRIOR FILING DATE: 1990-07-13
18 <150> PRIOR APPLICATION NUMBER: US 07/498,809
19 <151> PRIOR FILING DATE: 1990-03-23
21 <150> PRIOR APPLICATION NUMBER: US 07/379,076
22 <151> PRIOR FILING DATE: 1989-07-13
24 <150> PRIOR APPLICATION NUMBER: US 07/160,416
25 <151> PRIOR FILING DATE: 1988-02-25
27 <160> NUMBER OF SEQ ID NOS: 37
29 <170> SOFTWARE: PatentIn Ver. 2.0
31 <210> SEQ ID NO: 1
32 <211> LENGTH: 2932
33 <212> TYPE: DNA
34 <213> ORGANISM: Artificial Sequence
36 <220> FEATURE:
37 <223> OTHER INFORMATION: Description of Artificial Sequence: Nucleotide
38       sequence of expression vector, piH3
40 <400> SEQUENCE: 1
41 ggcgtaatct gctgcttgca aacaaaaaaa ccaccgctac cagcggtggt ttgtttgccg 60
43 gatcaagagc taccaactct ttttccgaag gaactggctt cagcagagcg cagataccaa 120
45 atactgtcct tctagtgtag ccgtagttag gccaccactt caagaactct gtagcaccgc 180
47 ctacatacct cgctctgcta atcctgttac cagtggctgc tgccagtggc gataagtcgt 240
49 gtcttaccgg gttggactca agacgatagt taccggataa ggcgcagcgg tcgggctgaa 300
51 cggggggttc gtgcacacag cccagcttgg agcgaacgac ctacaccgaa ctgagatacc 360
53 tacagcgtga gctatgagaa agcgccacgc ttcccgaagg agaaaggcg gacaggtatc 420
55 cggtaagcgg cagggtcgga acaggagagc gcacgaggga gcttccaggg ggaaacgcct 480
57 ggtatcttta tagtcctgtc gggtttcgcc acctctgact tgagcgtcga tttttgtgat 540
59 gctcgtcagg ggggcggagc ctatggaaaa acgccagcaa cgccgaatta ccgcggtgtt 600
61 tctcaacgta cactttaca gcggcgcgtc atttgatatg atgcgccccg cttcccgata 660
63 agggagcagg ccagtaaaag cattaccgt ggtggggttc ccgagcggcc aaagggagca 720
65 gactctaaat ctgccgtcat cgacttcgaa ggttcgaatc cttcccccac caccatcact 780
67 ttcaaaagtc cgaaagaatc tgctccctgc ttgtgtgttg gaggtcgctg agtagtgcgc 840
69 gagtaaaatt taagctacaa caaggcaagg cttgaccgac aattgcatga agaatctgct 900
71 tagggttagg cgttttgcgc tgcttcgcga tgtacgggcc agatatacgc gttgacattg 960
73 attattgact agttattaat agtaatcaat tacggggtca ttagttcata gcccatatat 1020
75 ggagttccgc gttacataac ttacggtaaa tggcccgcct ggctgaccgc ccaacgaccc 1080
77 ccgcccattg acgtcaataa tgacgtatgt tcccatagta acgccaatag ggactttcca 1140
79 ttgacgtcaa tgggtggact atttacggta aactgcccac ttggcagtac atcaagtgta 1200
81 tcatatgcca agtacgcccc ctattgacgt caatgacggt aaatggcccg cctggcatta 1260
83 tgcccagtac atgaccttat gggactttcc tacttggcag tacatctacg tattagtcat 1320
```

```
 85 cgctattacc atggtgatgc ggttttggca gtacatcaat gggcgtggat agcggtttga 1380
 87 ctcacgggga tttccaagtc tccaccccat tgacgtcaat gggagtttgt tttggcacca 1440
 89 aaatcaacgg gactttccaa aatgtcgtaa caactccgcc ccattgacgc aaatgggcgg 1500
 91 aattcctggg cgggactggg gagtggcgag ccctcagatg ctgcatataa gcagctgctt 1560
 93 tttgcctgta ctgggtctct ctggttagac cagatctgag cctgggagct ctctggctaa 1620
 95 ctagagaacc cactgcttaa gcctcaataa agcttctaga gatccctcga cctcgaggga 1680
 97 tcttccatac ctaccagttc tgcgcctgca ggtcgcggcc gcgactctag aggatctttg 1740
 99 tgaaggaacc ttacttctgt ggtgtgacat aattggacaa actacctaca gagatttaaa 1800
101 gctctaaggt aaatataaaa tttttaagtg tataatgtgt taaactactg attctaattg 1860
103 tttgtgtatt ttagattcca acctatggaa ctgatgaatg ggagcagtgg tggaatgcct 1920
105 ttaatgagga aaacctgttt tgctcagaag aaatgccatc tagtgatgat gaggctactg 1980
107 ctgactctca acattctact cctccaaaaa agaagagaaa ggtagaagac cccaaggact 2040
109 ttccttcaga attgctaagt ttttttgagtc atgctgtgtt tagtaataga actcttgctt 2100
111 gctttgctat ttacaccaca aaggaaaaag ctgcactgct atacaagaaa attatggaaa 2160
113 aatattctgt aacctttata agtaggcata acagttataa tcataacata ctgttttttc 2220
115 ttactccaca caggcataga gtgtctgcta ttaataacta tgctcaaaaa ttgtgtacct 2280
117 ttagcttttt aatttgtaaa ggggttaata aggaatattt gatgtatagt gccttgacta 2340
119 gagatcataa tcagccatac cacatttgta gaggttttac ttgctttaaa aaacctccca 2400
121 cacctccccc tgaacctgaa acataaaatg aatgcaattg ttgttgttaa cttgtttatt 2460
123 gcagcttata atggttacaa ataaagcaat agcatcacaa atttcacaaa taaagcattt 2520
125 ttttcactgc attctagttg tggtttgtcc aaactcatca atgtatctta tcatgtctgg 2580
127 atcctgtgga atgtgtgtca gttagggtgt ggaaagtccc caggctcccc agcaggcaga 2640
129 agtatgcaaa gcatgcatct caattagtca gcaaccaggt gtggaaagtc cccaggctcc 2700
131 ccagcaggca gaagtatgca aagcatgcat ctcaattagt cagcaaccat agtcccgccc 2760
133 ctaactccgc ccatcccgcc cctaactccg cccagttccg cccattctcc gccccatggc 2820
135 tgactaattt tttttattta tgcagaggcc gaggccgcct cggcctctga gctattccag 2880
137 aagtagtgag gaggcttttt tggaggccta ggcttttgca aaaagctaat tc         2932
```

```
140 <210> SEQ ID NO: 2
141 <211> LENGTH: 1504
142 <212> TYPE: DNA
143 <213> ORGANISM: Homo sapiens
145 <220> FEATURE:
146 <221> NAME/KEY: CDS
147 <222> LOCATION: (7)..(1059)
149 <400> SEQUENCE: 2
150 cctaag atg agc ttt cca tgt aaa ttt gta gcc agc ttc ctt ctg att       48
151        Met Ser Phe Pro Cys Lys Phe Val Ala Ser Phe Leu Leu Ile
152         1               5                  10
154 ttc aat gtt tct tcc aaa ggt gca gtc tcc aaa gag att acg aat gcc      96
155 Phe Asn Val Ser Ser Lys Gly Ala Val Ser Lys Glu Ile Thr Asn Ala
156  15                  20                  25                  30
158 ttg gaa acc tgg ggt gcc ttg ggt cag gac atc aac ttg gac att cct     144
159 Leu Glu Thr Trp Gly Ala Leu Gly Gln Asp Ile Asn Leu Asp Ile Pro
160                     35                  40                  45
162 agt ttt caa atg agt gat gat att gac gat ata aaa tgg gaa aaa act     192
163 Ser Phe Gln Met Ser Asp Asp Ile Asp Asp Ile Lys Trp Glu Lys Thr
164                 50                  55                  60
166 tca gac aag aaa aag att gca caa ttc aga aaa gag aaa gag act ttc     240
167 Ser Asp Lys Lys Lys Ile Ala Gln Phe Arg Lys Glu Lys Glu Thr Phe
```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/836,544A

DATE: 12/06/2001
TIME: 12:15:40

Input Set : A:\11-881.app
Output Set: N:\CRF3\12062001\I836544A.raw

```
168         65                    70                    75
170 aag gaa aaa gat aca tat aag cta ttt aaa aat gga act ctg aaa att   288
171 Lys Glu Lys Asp Thr Tyr Lys Leu Phe Lys Asn Gly Thr Leu Lys Ile
172         80                    85                    90
174 aag cat ctg aag acc gat gat cag gat atc tac aag gta tca ata tat   336
175 Lys His Leu Lys Thr Asp Asp Gln Asp Ile Tyr Lys Val Ser Ile Tyr
176  95                   100                   105                   110
178 gat aca aaa gga aaa aat gtg ttg gaa aaa ata ttt gat ttg aag att   384
179 Asp Thr Lys Gly Lys Asn Val Leu Glu Lys Ile Phe Asp Leu Lys Ile
180                   115                   120                   125
182 caa gag agg gtc tca aaa cca aag atc tcc tgg act tgt atc aac aca   432
183 Gln Glu Arg Val Ser Lys Pro Lys Ile Ser Trp Thr Cys Ile Asn Thr
184                   130                   135                   140
186 acc ctg acc tgt gag gta atg aat gga act gac ccc gaa tta aac ctg   480
187 Thr Leu Thr Cys Glu Val Met Asn Gly Thr Asp Pro Glu Leu Asn Leu
188         145                   150                   155
190 tat caa gat ggg aaa cat cta aaa ctt tct cag agg gtc atc aca cac   528
191 Tyr Gln Asp Gly Lys His Leu Lys Leu Ser Gln Arg Val Ile Thr His
192         160                   165                   170
194 aag tgg acc acc agc ctg agt gca aaa ttc aag tgc aca gca ggg aac   576
195 Lys Trp Thr Thr Ser Leu Ser Ala Lys Phe Lys Cys Thr Ala Gly Asn
196 175                   180                   185                   190
198 aaa gtc agc aag gaa tcc agt gtc gag cct gtc agc tgt cca gag aaa   624
199 Lys Val Ser Lys Glu Ser Ser Val Glu Pro Val Ser Cys Pro Glu Lys
200                   195                   200                   205
202 ggt ctg gac atc tat ctc atc att ggc ata tgt gga gga ggc agc ctc   672
203 Gly Leu Asp Ile Tyr Leu Ile Ile Gly Ile Cys Gly Gly Gly Ser Leu
204                   210                   215                   220
206 ttg atg gtc ttt gtg gca ctg ctc gtt ttc tat atc acc aaa agg aaa   720
207 Leu Met Val Phe Val Ala Leu Leu Val Phe Tyr Ile Thr Lys Arg Lys
208         225                   230                   235
210 aaa cag agg agt cgg aga aat gat gag gag ctg gag aca aga gcc cac   768
211 Lys Gln Arg Ser Arg Arg Asn Asp Glu Glu Leu Glu Thr Arg Ala His
212         240                   245                   250
214 aga gta gct act gaa gaa agg ggc cgg aag ccc cac caa att cca gct   816
215 Arg Val Ala Thr Glu Glu Arg Gly Arg Lys Pro His Gln Ile Pro Ala
216 255                   260                   265                   270
218 tca acc cct cag aat cca gca act tcc caa cat cct cct cca cca cct   864
219 Ser Thr Pro Gln Asn Pro Ala Thr Ser Gln His Pro Pro Pro Pro Pro
220                   275                   280                   285
222 ggt cat cgt tcc cag gca cct agt cat cgt ccc ccg cct cct gga cac   912
223 Gly His Arg Ser Gln Ala Pro Ser His Arg Pro Pro Pro Pro Gly His
224                   290                   295                   300
226 cgt gtt cag cac cag cct cag aag agg cct cct gct ccg tcg ggc aca   960
227 Arg Val Gln His Gln Pro Gln Lys Arg Pro Pro Ala Pro Ser Gly Thr
228         305                   310                   315
230 caa gtt cac cag cag aaa ggc ccg ccc ctc ccc aga cct cga gtt cag  1008
231 Gln Val His Gln Gln Lys Gly Pro Pro Leu Pro Arg Pro Arg Val Gln
232         320                   325                   330
```

```
234 cca aaa cct ccc cat ggg gca gca gaa aac tca ttg tcc cct tcc tct   1056
235 Pro Lys Pro Pro His Gly Ala Ala Glu Asn Ser Leu Ser Pro Ser Ser
236 335               340               345               350
238 aat taaaaaagat agaaactgtc tttttcaata aaaagcactg tggatttctg        1109
239 Asn
241 ccctcctgat gtgcatatcc gtacttccat gaggtgtttt ctgtgtgcag aacattgtca 1169
243 cctcctgagg ctgtgggcca cagccacctc tgcatcttcg aactcagcca tgtggtcaac 1229
245 atctggagtt tttggtctcc tcagagagct ccatcacacc agtaaggaga agcaatataa 1289
247 gtgtgattgc aagaatggta gaggaccgag cacagaaatc ttagagattt cttgtcccct 1349
249 ctcaggtcat gtgtagatgc gataaatcaa gtgattggtg tgcctgggtc tcactacaag 1409
251 cagcctatct gcttaagaga ctctggagtt tcttatgtgc cctggtggac acttgcccac 1469
253 catcctgtga gtaaaagtga aataaaagct ttgac                            1504
256 <210> SEQ ID NO: 3
257 <211> LENGTH: 351
258 <212> TYPE: PRT
259 <213> ORGANISM: Homo sapiens
261 <400> SEQUENCE: 3
262 Met Ser Phe Pro Cys Lys Phe Val Ala Ser Phe Leu Leu Ile Phe Asn
263   1               5               10              15
265 Val Ser Ser Lys Gly Ala Val Ser Lys Glu Ile Thr Asn Ala Leu Glu
266           20              25              30
268 Thr Trp Gly Ala Leu Gly Gln Asp Ile Asn Leu Asp Ile Pro Ser Phe
269           35              40              45
271 Gln Met Ser Asp Asp Ile Asp Asp Ile Lys Trp Glu Lys Thr Ser Asp
272       50              55              60
274 Lys Lys Lys Ile Ala Gln Phe Arg Lys Glu Lys Glu Thr Phe Lys Glu
275   65              70              75              80
277 Lys Asp Thr Tyr Lys Leu Phe Lys Asn Gly Thr Leu Lys Ile Lys His
278               85              90              95
280 Leu Lys Thr Asp Asp Gln Asp Ile Tyr Lys Val Ser Ile Tyr Asp Thr
281       '       100             105             110
283 Lys Gly Lys Asn Val Leu Glu Lys Ile Phe Asp Leu Lys Ile Gln Glu
284           115             120             125
286 Arg Val Ser Lys Pro Lys Ile Ser Trp Thr Cys Ile Asn Thr Thr Leu
287       130             135             140
289 Thr Cys Glu Val Met Asn Gly Thr Asp Pro Glu Leu Asn Leu Tyr Gln
290 145               150             155             160
292 Asp Gly Lys His Leu Lys Leu Ser Gln Arg Val Ile Thr His Lys Trp
293               165             170             175
295 Thr Thr Ser Leu Ser Ala Lys Phe Lys Cys Thr Ala Gly Asn Lys Val
296           180             185             190
298 Ser Lys Glu Ser Ser Val Glu Pro Val Ser Cys Pro Glu Lys Gly Leu
299           195             200             205
301 Asp Ile Tyr Leu Ile Ile Gly Ile Cys Gly Gly Gly Ser Leu Leu Met
302       210             215             220
304 Val Phe Val Ala Leu Leu Val Phe Tyr Ile Thr Lys Arg Lys Lys Gln
305 225               230             235             240
307 Arg Ser Arg Arg Asn Asp Glu Glu Leu Glu Thr Arg Ala His Arg Val
308               245             250             255
```

RAW SEQUENCE LISTING   DATE:. 12/06/2001
PATENT APPLICATION: US/09/836,544A   TIME: 12:15:40

Input Set : A:\11-881.app
Output Set: N:\CRF3\12062001\I836544A.raw

```
310 Ala Thr Glu Glu Arg Gly Arg Lys Pro His Gln Ile Pro Ala Ser Thr
311         260             265             270
313 Pro Gln Asn Pro Ala Thr Ser Gln His Pro Pro Pro Pro Gly His
314         275             280             285
316 Arg Ser Gln Ala Pro Ser His Arg Pro Pro Pro Gly His Arg Val
317     290             295             300
319 Gln His Gln Pro Gln Lys Arg Pro Pro Ala Pro Ser Gly Thr Gln Val
320 305             310             315             320
322 His Gln Gln Lys Gly Pro Pro Leu Pro Arg Pro Arg Val Gln Pro Lys
323             325             330             335
325 Pro Pro His Gly Ala Ala Glu Asn Ser Leu Ser Pro Ser Ser Asn
326         340             345             350
329 <210> SEQ ID NO: 4
330 <211> LENGTH: 874
331 <212> TYPE: DNA
332 <213> ORGANISM: Homo sapiens
334 <220> FEATURE:
335 <221> NAME/KEY: CDS
336 <222> LOCATION: (13)..(723)
338 <400> SEQUENCE: 4
339 gcccgacgag cc atg gtt gct ggg agc gac gcg ggg cgg gcc ctg ggg gtc  51
340                Met Val Ala Gly Ser Asp Ala Gly Arg Ala Leu Gly Val
341                 1             5             10
343 ctc agc gtg gtc tgc ctg ctg cac tgc ttt ggt ttc atc agc tgt ttt  99
344 Leu Ser Val Val Cys Leu Leu His Cys Phe Gly Phe Ile Ser Cys Phe
345         15             20             25
347 tcc caa caa ata tat ggt gtt gtg tat ggg aat gta act ttc cat gta 147
348 Ser Gln Gln Ile Tyr Gly Val Val Tyr Gly Asn Val Thr Phe His Val
349 30             35             40             45
351 cca agc aat gtg cct tta aaa gag gtc cta tgg aaa aaa caa aag gat 195
352 Pro Ser Asn Val Pro Leu Lys Glu Val Leu Trp Lys Lys Gln Lys Asp
353             50             55             60
355 aaa gtt gca gaa ctg gaa aat tct gaa ttc aga gct ttc tca tct ttt 243
356 Lys Val Ala Glu Leu Glu Asn Ser Glu Phe Arg Ala Phe Ser Ser Phe
357             65             70             75
359 aaa aat agg gtt tat tta gac act gtg tca ggt agc ctc act atc tac 291
360 Lys Asn Arg Val Tyr Leu Asp Thr Val Ser Gly Ser Leu Thr Ile Tyr
361             80             85             90
363 aac tta aca tca tca gat gaa gat gag tat gaa atg gaa tcg cca aat 339
364 Asn Leu Thr Ser Ser Asp Glu Asp Glu Tyr Glu Met Glu Ser Pro Asn
365         95             100             105
367 att act gat acc atg aag ttc ttt ctt tat gtg ctt gag tct ctt cca 387
368 Ile Thr Asp Thr Met Lys Phe Phe Leu Tyr Val Leu Glu Ser Leu Pro
369 110             115             120             125
371 tct ccc aca cta act tgt gca ttg act aat gga agc att gaa gtc caa 435
372 Ser Pro Thr Leu Thr Cys Ala Leu Thr Asn Gly Ser Ile Glu Val Gln
373             130             135             140
375 tgc atg ata cca gag cat tac aac agc cat cga gga ctt ata atg tac 483
376 Cys Met Ile Pro Glu His Tyr Asn Ser His Arg Gly Leu Ile Met Tyr
```

VERIFICATION SUMMARY   DATE: 12/06/2001
PATENT APPLICATION: US/09/836,544A   TIME: 12:15:41

Input Set : A:\11-881.app
Output Set: N:\CRF3\12062001\I836544A.raw

OIPE

RAW SEQUENCE LISTING
PATENT APPLICATION: **US/09/836,544**

DATE: 12/06/2001
TIME: 10:11:02

Input Set : **A:\11-88L.txt**
Output Set: **N:\CRF3\12062001\I836544.raw**

Does Not Comply
Corrected Diskette Needed

```
 3 <110> APPLICANT: The General Hospital Corporation
 5 <120> TITLE OF INVENTION: Rapid Immunoselection Cloning Method
 7 <130> FILE REFERENCE: 11-88L
 9 <140> CURRENT APPLICATION NUMBER: US 09/836,544
10 <141> CURRENT FILING DATE: 2001-04-17
12 <150> PRIOR APPLICATION NUMBER: US 07/983,647
13 <151> PRIOR FILING DATE: 1992-12-01
15 <150> PRIOR APPLICATION NUMBER: US 07/553,759
16 <151> PRIOR FILING DATE: 1990-07-13
18 <150> PRIOR APPLICATION NUMBER: US 07/498,809
19 <151> PRIOR FILING DATE: 1990-03-23
21 <150> PRIOR APPLICATION NUMBER: US 07/379,076
22 <151> PRIOR FILING DATE: 1989-07-13
24 <150> PRIOR APPLICATION NUMBER: US 07/160,416
25 <151> PRIOR FILING DATE: 1988-02-25
27 <160> NUMBER OF SEQ ID NOS: 37
29 <170> SOFTWARE: PatentIn Ver. 2.0
```

ERRORED SEQUENCES

```
3446 <210> SEQ ID NO: 37
3447 <211> LENGTH: 12
3448 <212> TYPE: DNA
3449 <213> ORGANISM: Artificial Sequence
3451 <220> FEATURE:
3452 <223> OTHER INFORMATION: Description of Artificial Sequence:
3453       Oligonucleotide primer
3455 <400> SEQUENCE: 37
3456 ctttagagca ca                                                   12
```

E--> 3460 (64)

VERIFICATION SUMMARY
PATENT APPLICATION: US/09/836,544

DATE: 12/06/2001
TIME: 10:11:04

Input Set : **A:\11-88L.txt**
Output Set: **N:\CRF3\12062001\I836544.raw**

L:3460 M:254 E: No. of Bases conflict, LENGTH:Input:64 Counted:12 SEQ:37